CDF note 10104

Search for Light Higgs Boson from Top Quark Decays

The CDF Collaboration

May 3, 2010

Abstract

We present a search for light (with mass below 2m$_b$) NMSSM pseudo-scalar Higgs boson A originating from top quark decays: $t \to H^{\pm}b \to W^{\pm(*)}A\,b$. Under assumption that A decays into $\tau\bar{\tau}$, we search in events with lepton + $\geq$ 3 jets and $\geq$ 1 b-tag for an additional isolated track in the p_T range from 3 to 20 GeV/c. Using a fit to the isolated track p_T spectrum we derive limits on the branching ratio of $t \to H^{\pm}b$ for various $H^{\pm}$ and A masses.

1 Introduction

The minimal supersymmetric model (MSSM) extends the Higgs sector of the standard model (SM) by including two Higgs doublet fields thats leads to five Higgs bosons: a light and a heavy SM-like CP-even Higgs (denoted h and H respectively), a CP-odd Higgs (A), and a pair of charged higgses ($H^{\pm}$). The MSSM Higgs sector is constrained by the decay mode independent LEP limits on SM-like Higgs bosons: $m_h > 82$ GeV, which is in conflict with theoretical calculations that require m_h to be less than 82 GeV[1]. However, if the Higgs sector of MSSM is extended with an additional singlet chiral superfield, m_h can theoretically be above current experimental limits. The addition of the singlet chiral superfield to the MSSM results in the next-to-minimal supersymmetric model (NMSSM), which contains an additional CP-even and CP-odd neutral Higgs bosons, as well as an additional neutralino, as compared to the MSSM. Masses of $m_A > 2m_b$ have been mostly ruled out by the LEP experiments, but for $m_A < 2m_b$ there has yet to be any limits placed (although there are some constraints [2, 3]).

For a large region of parameter space (especially for $\tan\beta \lesssim 2.5$) $H^{\pm}$ would most likely be within a reach of the Tevatron, with $m_{H^{\pm}}$ having a little dependence on other masses of SUSY particles. In the scenario of $m_A < 2m_b$, LEP limits on $H^{\pm}$ generally don't apply, since LEP searches are focused on $e^+e^- \to H^+H^-$, with $H^+ \to c\bar{s}$, $H^+ \to \tau^+\nu_\tau$, or $H^+ \to W^{+(*)}A$ assuming $m_A > 2m_b$. If m_A is less than $2m_b$, then the branching ratio of $H^+ \to W^{+(*)}A$ is always greater than 0.5 for $m_{H^{\pm}} \approx 100$ GeV/c^2, thus making the LEP limits not applicable. Figure 1 shows that in the case of $m_A < 2m_b$ the dominant decay mode of A is into $\tau\bar{\tau}$ [4].

An additional motivation for this search comes from the observed 2.8 σ deviation of lepton universality in the decay of W bosons at LEP[5]:

$\mathcal{B}(W^{\pm} \to \tau\nu) / \mathcal{B}(W^{\pm} \to e\nu) = 1.070 \pm 0.029$ and

$\mathcal{B}(W^{\pm} \to \tau\nu) / \mathcal{B}(W^{\pm} \to \mu\nu) = 1.076 \pm 0.028$.

It is argued, that this could be due to H^+H^- production with subsequent $H^{\pm} \to \tau\nu$ decays, in case of $m_{H^{\pm}} \approx m_{W^{\pm}}$[6]. Again, this would imply that $H^{\pm}$ could be within the Tevatron reach, and could be observed in top decays, if $\mathcal{B}(t \to H^{\pm}b)$ is sufficiently large.

In this analysis we search for decays $t \to H^{\pm}b \to W^{\pm(*)}A\,b$. To avoid existing experimental and theoretical limits we assume $m_{W^{\pm}} \lesssim m_{H^{\pm}} < m_t - m_b$, and $m_A < 2m_b$. Theoretically it is expected that if $m_{W^{\pm}} \approx m_{H^{\pm}}$ then the branching ratio of $t \to H^+b$ is approximately 0.4 to 0.1 for $\tan\beta$ in the range of 1 to 2.5, which is fairly independent of superpartner masses and m_A, given that $m_A < 2m_b$ [6].

There have been a number of searches for $H^{\pm}$ in top decays at CDF and D0 that have set limits with assumptions that top can decay as $t \to H^{\pm}b$ with the $H^{\pm}$ decaying as: $H^+ \to \tau\bar{\nu}$, $H^+ \to c\bar{s}$, $H^+ \to t^*b \to W^+b\bar{b}$, and $H^+ \to W^+A \to W^+b\bar{b}$ [7, 8], however none have considered the possibility of $m_A < 2m_b$, thus leaving this region of interest unconstrained.

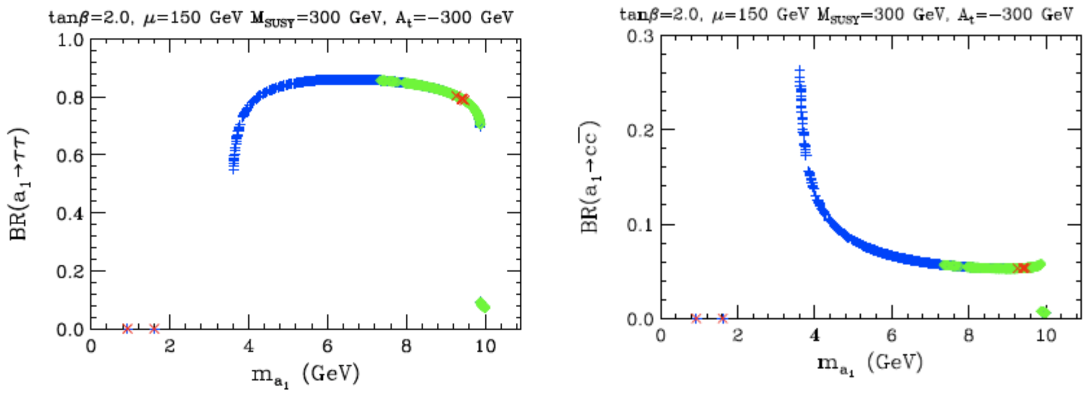

Figure 1: $\mathcal{B}\,(A \to \tau\bar{\tau})$ (left) and $\mathcal{B}\,(A \to c\bar{c})$ (right) vs. m_A for $\tan\beta = 2$, and a particular choice of SUSY masses. Figures are from Ref. [4].

2 Method Overview

The main difference between $t \to W^+b$ and $t \to H^+b \to W^{+(*)}A\,b$ events is the presence of two τ's from the A in the final state. These τ 's so far could have easily gone undetected, since they typically have low p_T. To improve the τ ID efficiency we identify τ's as isolated tracks in the p_T range between 3 and 20 GeV/c.

Such isolated tracks can originate from a number of sources: low-p_T $e^\pm$ and $\mu^\pm$ leptons, $\tau^\pm$ leptons, hadrons with only one track reconstructed within a jet cone, charged particles ($\pi^\pm$, $K^\pm$, etc.). These tracks can be produced from both the hard interaction that produces high- p_T electron or muon candidate responsible for triggering the event as well as from additional soft parton interactions. We refer to the latter as the "Underlying Event" (UE).

To study the modeling of UE tracks we use control regions: multi-jet QCD events, Z events and W events (a single lepton and less than three jets to reduce the contribution from $t\bar{t}$). We find that the modeling of UE tracks from Monte Carlo is in poor agreement with data in terms of both the rate, as well as the p_T-spectra of the tracks. Therefore we use data to model the characteristics of UE tracks. By comparing different data sources, we find the p_T spectra of isolated tracks from UE to be similar. Therefore to enhance statistics we model UE tracks using the data triggered in presence of jets. This is discussed in more detail in Sec. 4.

The pre-signal region is defined as events with one and only one lepton, three or more jets and at least one b-tag. The non-UE isolated tracks come from physics processes, where more than one lepton is produced, but only one is identified, such as for instance from Z/γ^* events where one lepton triggers the event, while the other one has a p_T less than 20 GeV/c, or a $\tau^\pm$ that only leaves a low p_T track. The tracks from W and Z vector bosons have usually harder p_T spectra than those of τ's from A decays.

We search for presence of A bosons in top decays by requiring at least one low-p_T isolated track (in addition to pre-signal event selection criteria) and performing the kinematic fit to the isolated track p_T spectrum of a combination of UE and non-UE SM sources and a new physics signal. In case when more than one isolated track is found in an event, the highest p_T track is being used.

3 Standard Model and Signal Modeling

We use PYTHIA for modeling $t\bar{t}$ events (assuming $m_t = 172.5$ GeV/c^2) and ALPGEN+PYTHIA for modeling W+X and Z/γ^*+X. The ALPGEN MC samples with different parton multiplicities are merged according to their relative leading order cross sections, and the overlap between the light and heavy flavor samples is removed using the procedure developed for the SECVTX top cross section analysis [9].

Signal MC is produced using PYTHIA generator with the anti-top quark always decaying as $\bar{t} \to W^- b$, while the top quark is forced to always decay as $t \to H^+ b \to W^{+^{(*)}} A\, b$. The $W^\pm$ bosons are allowed to decay to all SM final states. The A particle is only allowed to decay into $\tau\bar{\tau}$. The MC samples with both tops decaying into $H^\pm b$ used for cross-checks.

We use an analytic method to correct the signal for varying $\mathcal{B}\left(t \to H^\pm b\right)$. Kinematic acceptance due to the modified decay chain of signal events is linearly dependent on the signal branching ratio, e.g. if the fractional acceptance difference between SM $t\bar{t}$ and the signal is $\Delta^A_{0.5}$, and the signal branching ratio is $\mathcal{B}$, then the signal acceptance is $\mathcal{A}\left(\mathcal{B}\right) = \mathcal{A}_{SM}\left(1.0 + 2\mathcal{B}\Delta^A_{0.5}\right)$.

The signal MC is also corrected for the fraction of events with two A's and > 1 track according to the equation: $(1 - \prod\left(1 - \mathcal{B} * \epsilon\right)) \times (1 - 0.033\mathcal{B}) = (2\mathcal{B} * \epsilon - \mathcal{B}^2\epsilon^2) \times (1 - 0.033\mathcal{B})$, where ϵ is the efficency of an A to produce at least one track passing the identification requirements. The factor $(1 - 0.033\mathcal{B})$ accounts for the possibility a tau from the second A in the event be near a tau from from the first A and violate the isolation requirements. The value of 0.033 is derived from a signal MC sample with both tops decaying to $H^\pm b$.

Prior to the isolated track requirement we employ the same event selection criteria as in the top cross section measurement [9]. The $t\bar{t}$ cross section is measured from the data using the same method as in [9]. This allows us to avoid experimental uncertainties on $t\bar{t}$ normalization associated with luminosity, b-tagging efficiency jet energy scale and etc. Assuming $m_{top} = 172.5$ GeV/c^2 we measure a $t\bar{t}$ cross section of $\sigma_{t\bar{t}} = 7.78 \pm 0.39_{stat} \pm 0.54_{syst}$. The event yields from different physics sources are listed in Table 1.

We select tracks with $p_T > 3$ GeV/c in the fiducial detector region $|\eta| < 1.1$, originated near the reconstructed primary vertex of the event and having additional standard track quality requirements, such as a sufficient number of reconstructed hits in the tracking chamber and etc. The tracks are also required to be not within $\Delta R < 0.4$

cone of the reconstructed lepton (e or μ) or a jet and have a track-based isolation:

$$TrkIsol = \frac{p_T\,(candidate)}{p_T\,(candidate) + \sum p_T^{trk}} > 0.9$$

where $\sum p_T^{trk}$ is the sum of the p_T of tracks within a cone of 0.4 of the candidate track, that have $p_T > 0.5$ GeV/c, $\Delta Z\,(trk, candidate) < 5$ cm, and pass the track quality requirements.

4 Underlying Event Modeling

The tracks from the UE constitute a major background. We model the UE track p_T spectrum using data. We find that the p_T spectrum of isolated tracks is consistent between many data sources. We study dilepton events by selecting Z-candidates and looking for an additional isolated track. We study QCD multi-jet events using data triggered on presence of jets. We also study lepton + jets by selecting events with only one or two jets, since three and more jets events correspond to our signal region.

The isolated track p_T spectra for pretag (no b-tagging requirement) lepton + jets, dilepton and multi-jet events are shown in Figure 2 for both one and two jet bins. The track p_T spectrum for lepton+jets events needs to be corrected for tracks from vector bosons from Z/γ^*, diboson, and top events. This is done by subtracting tracks originating from a $W^\pm$ or Z/γ^* using MC. We trace the reconstructed track back to originating from a charged particle in the decay chain of the boson's daughter particles,

Event Yields Predictions for Signal Region ≥ 3 Jets		
Source	Pretag	Tagged
$t\bar{t}$	1347.1	817.9
W+LF	1140.4	49.2
W+cj	142.8	16.0
$W + c\bar{c}$	242.9	34.2
$W + b\bar{b}$	146.8	67.9
Diboson	118.7	11.2
Z+LF	179.7	7.2
Z+HF	25.0	11.6
Non-W	549.8	45.5
Total	3893.00	1060.6
Data	3893	1052

Table 1: The signal region predicted event yields for 2.7 fb^{-1} of data from the different physics sources in the ≥ 3 jet bin, for $\sigma_{t\bar{t}} = 7.78$ pb.

and normalize backgrounds accordingly. As seen in the Figure after the correction the p_T spectra agree.

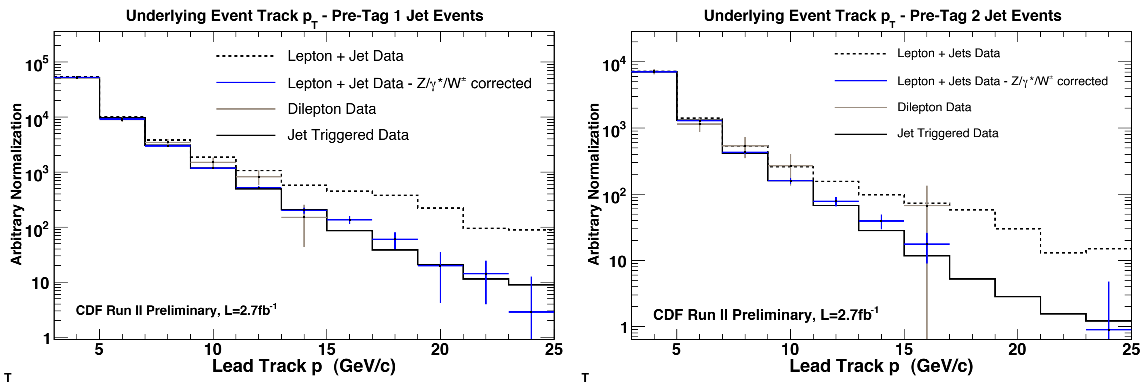

Figure 2: The isolated track p_T spectrum for pretag lepton+jets, dilepton events with one jet (left) and two jets (right). The track $p_T$spectrum for lepton+jets needs to be corrected for tracks from vector bosons from Z/γ^*, diboson, and top events before a proper comparison. Both raw and corrected spectra are shown. The multi-jet sample's track p_T spectrum is from events with two jets.

We performed many additional cross-checks, such as comparing UE track p_T spectra for different multi-jet samples, the dependence on the jet multiplicity, the number of primary vertices, presence of the b-tag in event, and observed no difference in the track p_T spectra, although the differences in the rate of events with at least one isolated track are somewhat larger than individual statistical uncertainties, ranging from 6.9% to 8.2%.

We find one sub-sample where the track p_T is a bit different and has a harder p_T spectrum. This is a mono-jet sample, i.e. events with one and only one reconstructed jet. Since these events correspond to the situation where at least one of the jets is not reconstructed (falls into the beam pipe region, fails the jet E_T requirement or falls into the detector crack region), we do not consider these events in forming the track p_T spectrum. The fact that mono-jet events produce a bias is also confirmed by $\Delta\phi$ distribution between the isolated track and the leading jet in the event. In monojet data there is a trend that the track and and the jet are back-to-back confirming that the bias is due to the additional tracks from the non-reconstructed jet. Two or higher jet multiplicity events do not show the same features and are consistent with each other.

The UE track p_T distribution from multi-jets data is used to model the UE contribution. In the final fit in the signal region we allow the rate of UE tracks to float freely. We use MC events only to model the track p_T spectra corresponding to vector boson daughters. We therefore require the MC tracks to match vector boson daughters. Next, since we use only one highest p_T track per event in the fit, we correct the MC-based track p_T spectra for the probability that a UE track could have a higher p_T track, and thus would be selected as a leading track in an event. To correct for that the MC-matched track p_T spectra are weighted according to: $w = 1 - \mathcal{R}_{UE}^{track} * Prob\left(p_T^{UE} > p_T^{track}\right)$, where $\mathcal{R}_{UE}^{track}$ is the average rate of events with an UE track, and $Prob\left(p_T^{UE} > p_T^{track}\right)$ is the probability that a UE track has a higher p_T than the MC-matched track.

In addition, since we generate signal events with only one A particle, we also correct the signal MC spectrum for the fact that there could be two A's present in an event and the p_T spectrum would be harder. This is done by applying the weight: $w = 1 - \mathcal{R}_{A}^{track} * \mathcal{B}_{t\to H^{\pm}b} * Prob\left(p_T^{signal} > p_T^{track}\right)$.

5 Z/γ^*Cross Section using Lepton +Jets Data

An important cross-check to ensure that our isolated track modeling is adequate is extracting the Z/γ^*cross section in lepton plus jets events. As described above, Z/γ^*events contribute to lepton plus jets region, when the second lepton from $Z/\gamma^*$is not identified, but passes our isolated track requirements.

We perform the fit to the isolated track p_T spectrum, which represents a standard log-likelihood fit with UE and Z/γ^* contributions completely unconstrained, and top/diboson contributions constrained to 20% of their predictions. The UE distribution is constructed from multi-jet data events. Other p_T spectra are obtained from MC samples with isolated tracks matched to daughter tracks from vector boson decays. The fit is performed in the range $3 \le p_T \le 20$ GeV/c separately for one jet and two jet bins. The results of the fit are presented in Figure 3. Tracks from Z/γ^*events are typically harder than those from the UE events. The extracted contribution from Z/γ^*events matches the expected contribution within the statistical uncertainties. This confirms that the modeling of the UE track p_T spectrum is adequate, and we can proceed with search for A -boson decays in the $t\bar{t}$-dominant signal region.

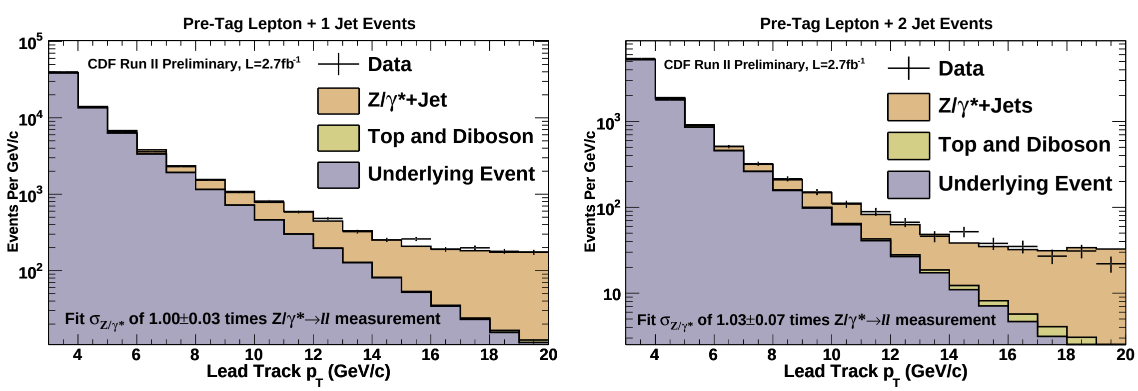

Figure 3: Fits for Z/γ^*+jets cross section in lepton + jets events data for ==1 jets (left), and ==2 jets (right). In both case the fit results are consistent with expected Z/γ^*contribution.

6 Limit Setting Procedure and Systematic Uncertainties

We employ the CL$_S$ limit setting procedure [10] to place limits on, or quote significance of excess in data. The CL$_S$ technique is based on the likelihood ratio test statistics:

$$Q = \frac{L(\text{data} \mid \text{signal} + \text{background})}{L(\text{data} \mid \text{background only})} \tag{1}$$

The test statistics Q_{obs} for data is compared to the values of Q obtained for a large number of pseudo-experiments generated under both the Standard Model hypothesis (H_0) and the Standard Model + New Physics Signal hypothesis (H_1). The quantity

$$CL_s = \frac{Prob_{H_1}(Q \leq Q_{obs})}{Prob_{H_0}(Q \leq Q_{obs})} \tag{2}$$

is used to set limits. In case $CL_s = 0.05$ the new physics signal is excluded at the 95% C.L., and $1 - Prob_{H_0}(Q \leq Q_{obs})$ defines a p-value for data under the Standard Model hypothesis, that quantifies a significance of an excess.

Systematic uncertainties are incorporated both in the likelihood fit and in generation of pseudo-experiments, and listed below.

7 Systematic Uncertainties

Because the UE contribution is allowed to float freely in the fit to data, there is no systematic uncertainty associated with the UE rate. For computing the expected limits, however, pseudo-experiments are generated and assumptions about the UE rate need to be made. We use the average rate across different data samples of 7.5%, and use the largest discrepancy in the rate between different data samples as a UE rate uncertainty in generation of pseudo-experiments. This corresponds to the relative rate uncertainty of 15%.

The largest variations in the UE isolated track p_T spectrum come from varying the H_T for events. Conservatively, we use the shapes obtained from multi-jet data for very low H_T and very high H_T as the UE shapes corresponding to +1 or -1 standard deviation shapes. Next, in the fit we allow the shape of the UE track p_T spectrum to continuously morph between these two extreme values.

As described in Sec. 3, the $t\bar{t}$ normalization is determined from data. The residual uncertainty comes from the systematic uncertainties in the $t\bar{t}$ cross section measurement [9]. Those include the uncertainties in the lepton ID, trigger and b-tagging efficiencies, the jet energy scale, as well as the uncertainty in the estimate of backgrounds to $t\bar{t}$ (9% total).

The Z/γ^*+ heavy flavor MC is normalized to data under the Z mass peak, with the dominant uncertainty due to limited statistics of $Z+$ tagged jet events in data (8%).

Events per 2.7 fb^{-1} in the signal region.

	SM Top	QCD & W	Diboson & Z/γ^*	UE	Signal	Data
Events	804.7	212.8	29.9	-	133.0	1052
Events with a Track	2.6±0.2	-	0.8±0.1	79.4±11.9	22.9±2.3	70

Table 2: Expected event yields before the track requirement (1st row), and with at least one isolated track (2nd row) with $\mathrm{p}_T > 3\,GeV/c$ in 2.7 fb^{-1}. Uncertainties are quoted on the number of events with tracks only. The signal corresponds to an example scenario with $m_{H^\pm} = 100\,\mathrm{GeV}/c^2$, $m_A = 8\,\mathrm{GeV}/c^2$, $\mathcal{R}(t \to H^\pm b) = 0.086$, excluded at 95% C.L.

The uncertainty in the diboson background is due to theory, luminosity and the jet energy scale (20%). However, its contribution to the signal region is very minor.

Since we require the isolated track not to be within a reconstructed jet, the systematic shift in the jet energy scale leads to events migrating to/from the signal region, which results in additional 3% uncertainty applied to all MC-based backgrounds.

8 Results

The expected event yields in the signal region are presented in Tab. 2. The first row in the table represents the numbers of expected and observed events before the isolated track requirement (the same as in Tab. 1). The second row shows the event yields after the isolated track requirement. Here the contribution in non-UE columns represent the expected numbers of events with the isolated tracks from the vector boson decays respectively. The quoted in the table expected event yield due the Underlying Event corresponds to the expected rate, while the actual normalization is obtained from the fit to the isolated track $p_T$spectrum in data, as it is shown in Fig. 4 (left). The column representing a number of signal events corresponds to an example scenario, excluded at 95% C.L.

As seen from Fig. 4 (left) no excess that can be attributed to considered non-SM top decays is observed. The data stands in a good agreement with predictions. The 95% C.L limits on the branching ratio of $t \to H^+ b$ as a function of $m_{H^\pm}$ and m_A are obtained and shown in Fig. 4 (right).

In conclusion, we performed a search for non-SM top decays $t \to H^\pm b \to W^{\pm(*)} A\, b$ within the NMSSM scenario. We see no evidence of τ's from light Higgs A decays, and set the first world's limits on the branching ratio of $t \to H^+ b$ in this mode.

References

[1] G. Abbiendi et al. [OPAL Collaboration], Eur. Phys. J. C **27**, 311 (2003) arXiv:hep-ex/0206022v1

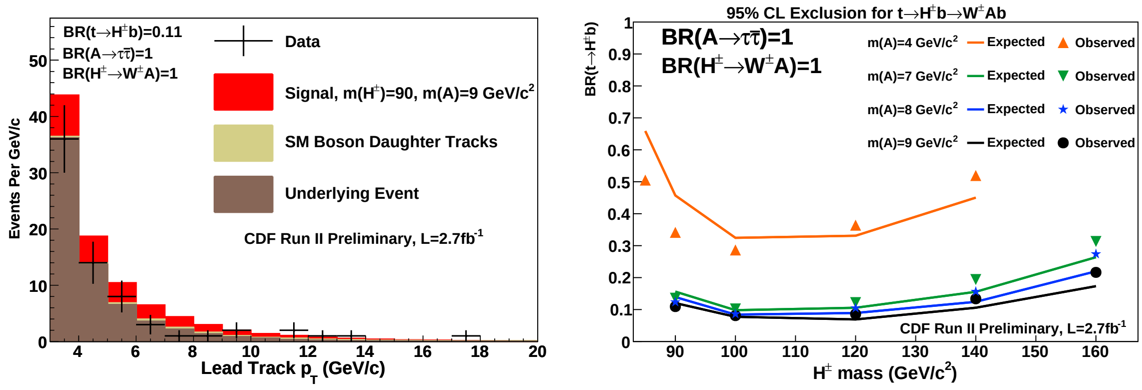

Figure 4: **LEFT:** The observed isolated track p_T spectrum. **RIGHT:** Expected and observed limits on the branching ratio of $t \to H^+b$.

[2] W. Love, et al.(CLEO), CLNS 08/2033, CLEO 08-16 to be submitted to PRL.

[3] G. Abbiendi et al. [OPAL Collaboration], Eur. Phys. J. C **23**, 397 (2002) arXiv:hep-ex/0111010

[4] R. Dermisek and J. Gunion, Phys. Rev. **D** 79 (2009) 055014.

[5] The LEP Collaborations, arXiv:hep-ex/0412015v2

[6] R. Dermisek, arXiv:hep-ph/0807.2135

[7] A. Abulencia et al. [CDF Collaboration], Phys. Rev. Lett. 96, 042003 (2006).

[8] T. Aaltonen et al., [CDF Collaboration], Phys. Rev. Lett. 103, 101803 (2009).

[9] T. Aaltonen et al., [CDF Collaboration], arXiv: hep-ex/1004.3224.

[10] T. Junk, Nucl. Instrum. Methods A 434 (1999).